Journal of Differential Equations **169**, 169–189 (2001)
doi:10.1006/jdeq.2000.3897, available online at http://www.idealibrary.com on IDEAL®

The Non-Holonomic Mechanics

Ivan Kupka

Department of Mathematics, Un. Pierre et Marie Curie-Paris VI, 4, Place Jussieu, 75230 Paris Cedex, France
E-mail: kupka@math.jussieu.fr

and

Waldyr M. Oliva

Department of Mathematics, Instituto Superior Técnico, Un. Técnica de Lisboa, Av. Rovisco Pais 1, Lisboa, Portugal
E-mail: wamoliva@math.ist.utl.pt

DEDICATED TO PROFESSOR JACK K. HALE ON THE OCCASION OF HIS 70TH BIRTHDAY

Two approaches for the study of mechanical systems with non-holonomic constraints are presented: d'Alembertian mechanics and variational (vakonomic) mechanics. The first one is equivalent to the d'Alembert principle and the second comes from a variational principle. In order to develop the two approaches, d'Alembertian and vakonomic trajectories are introduced. A generalized version of the classical Liouville theorem for the conservation of volume is proved in the context of d'Alembertian mechanics. A characterization for the notions of regular and singular curves is presented. The flow corresponding to the regular vakonomic trajectories is Hamiltonian. © 2001 Academic Press

1. INTRODUCTION

There are two well known approaches for the study of the classical non-holonomic mechanics: the so called *d'Alembertian non-holonomic mechanics* and the *variational non-holonomic* (*vakonomic*) *mechanics*. In both cases there is a *configuration space* represented by a connected C^∞ Riemannian manifold (M^n, g) and a (*non-holonomic*) *constraint* defined by a smooth (non-integrable) distribution $\mathscr{D} \subset TM$ with constant rank m, $0 < m < n$. The metric g, also denoted by $\langle\ ,\ \rangle$, defines the Levi–Civita connection and the *kinetic energy* K: $TM \to \mathbb{R}$ given by $K(\xi) = \frac{1}{2}\langle \xi, \xi\rangle$, $\xi \in TM$; the *potential energy* is a smooth function V: $M \to \mathbb{R}$ that will define the conservative field of external forces. The *trajectories* are absolutely continuous curves q: $t \in [a_0, a_1] \subset \mathbb{R} \to q(t) \in M$ $(q \in H^1(M, [a_0, a_1]))$, such that $\dot{q}(t) = \frac{Tq}{dt}(t)$

0022-0396/01 $35.00

$\in \mathscr{D}$ for a.e $t \in [a_0, a_1]$; one uses to say that the trajectories are curves in $H^1(M, [a_0, a_1])$ compatible with the distribution ($q \in H^1(M, \mathscr{D}, [a_0, a_1])$). In order to develop the two approaches, d'Alembertian and vakonomic trajectories are defined in Section 3 where we compare the two kinds of curves, properly.

In d'Alembertian non-holonomic mechanics, studied in Section 4, the trajectories satisfy the so called *d'Alembert principle* that states (see [C], [FO], [MP], [CF], [Ko]): the difference between the *acceleration* $\nabla_t \dot{q}$ of the trajectory $q = q(t)$ and the external force $(-grad\ V)(q(t))$ is orthogonal to $\mathscr{D}_{q(t)}$ for all $t \in [a_0, a_1]$ (here ∇_t is the covariant derivative of the Levi–Civita connection and *grad V* is defined by $dV(\cdot) = \langle grad\ V, \cdot \rangle$); the main results are Propositions 4.1 and 4.2 below where the first gives the equations of motion and the second is a generalized version, for d'Alembertian non-holonomic mechanics, of the famous Liouville theorem relative to the conservation of volume.

Section 5 deals with the structures of Hilbert manifolds of some special sets called $\mathscr{D}$*-spaces*, mainly $H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ (resp. $H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$) that is the set of all absolutely continuous curves $q\colon [a_0, a_1] \to M$, compatible with $\mathscr{D}$ such that $q(a_0) = m_0 \in M$ (resp. $q(a_0) = m_0$, $q(a_1) = m_1 \in M$) and it is also considered the corresponding evaluation map ev_1: $H^1(M, \mathscr{D}, [a_0, a_1], m_0) \to M$, $ev_1(q) = q(a_1)$. The regular and critical points of the smooth map ev_1 lying in $ev_1^{-1}(m_1) = H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$ are called *regular* and *singular curves*, respectively, associated to the value m_1 of ev_1. The singular curves are characterized in Proposition 5.2 and they do not depend on the Riemannian metric g (see Remark 5.1).

The variational non-holonomic (vakonomic) mechanics works with trajectories that are determined by a variational approach; in fact each vakonomic trajectory corresponding to the data $(M, K, \mathscr{D}, V)$ is a stationary point of a Lagrangian functional $\mathscr{L}$ given by $\mathscr{L}(q) = \int_{a_0}^{a_1} [\frac{1}{2} \|\dot{q}\|^2 - V(q)]\, dt$; $\mathscr{L}$ is defined in the Hilbert manifold $H^1(M, [a_0, a_1])$ and is restricted to the subset $H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$ where $q(a_0) = m_0$, $q(a_1) = m_1$ (see Section 3). The regular stationary points of $\mathscr{L}$ are the *vakonomic trajectories* that correspond to some presentations already considered, more recently, in the literature (see [AKN], [VG], [CF], and [Z]). The second order ordinary differential equation for the regular vakonomic trajectories is derived in Proposition 6.1 and it defines a flow of a Hamiltonian vector field on the manifold $TM = \mathscr{D} \times_M \mathscr{D}^{\perp}$ (see Proposition 6.2); so the solutions of that vector field are, then, of the type $(\dot{q}(t), P(t))$ where $\dot{q}(t) \in \mathscr{D}_{q(t)}$ and $P(t) \in \mathscr{D}^{\perp}_{q(t)}$, for all $t \in [a_0, a_1]$, $q = q(t)$ being a regular vakonomic trajectory. Locally, the components of $P(t)$ correspond to the classical Lagrange multipliers (see [AKN], [VG], [CF], [Z]).

For the sake of motivation and completeness we would like to mention that if we restrict ourselves to a free dynamics i.e., if the potential energy function V is zero, the non-holonomic mechanics is related with some geometric studies and concepts: d'Alembertian mechanics with the so called $\mathscr{D}$-geodesic flows and vakonomic mechanics with sub-Riemannian geometry. A study on $\mathscr{D}$-geodesic flows and the construction of examples of partially hyperbolic and Anosov flows, including an application to semi-simple Lie groups, is presented in [CKO]. For an exposition on sub-Riemannian geometry and its relation with other domains of mathematics, see [Ku]. Finally, we remark that some of the results of the present paper can be extended to more general Lagrangian functionals and also to non-linear constraints (see [AKN], [MP], [Ko], [VG], [CF],[Z]).

Before concluding this introduction, since Mechanics is not just an abstract mathematical theory but is relevant to many practical problems, it is only appropriate to ask the following question: does nature follow the d'Alembert or the vakonomic mechanics? A. D. Lewis and R. M. Murray, from CalTec, have performed careful experiments to address this question. They present their results in the paper [LM] and show that with the addition of friction terms to the d'Alembertian (non-holonomic in their terminology) model, there is a reasonable agreement between the experimental data and theoretical computations. The theoretical content of the paper [LM] is very much related to ours. Its setting is more "mechanical" and less coordinate free than ours. The constraints in [LM] are more general being "affine" instead of just linear as in the present paper. On the other hand the singular curves (Section 5.2 and Proposition 5.2), a typical occurence of the linear constraints, are not mentioned in [LM].

2. NOTATIONS AND BASIC DEFINITIONS

M will denote a smooth connected paracompact manifold and $g: TM \to \mathbb{R}_+$ a smooth Riemannian metric on M, also represented as $g(w) = \langle w, w \rangle$; TM (resp. T^*M) denotes the tangent (resp. cotangent) bundle of M and $\pi_{TM}: TM \to M$, $\pi_{T^*M}: T^*M \to M$ the associated projections. Smooth means C^∞ or real analytic (C^ω). ∇ will denote the Levi–Civita covariant differentiation operator associated to the Riemannian manifold (M, g). Finally, $K = \frac{1}{2} g$ will be called the kinetic energy of (M, g).

Definition 2.1. A mechanical system with non-holonomic constraints on (M, g) will be a quadruple $(M, K, \mathscr{D}, V)$ where $\mathscr{D}$ is a constant rank smooth non integrable distribution (that is, a vector subbundle of TM) and a smooth function $V: M \to \mathbb{R}$ called the potential energy. $\mathscr{D}$ represents the non-holonomic constraint.

$\mathscr{D}^{\perp}$ will denote the distribution orthogonal complement of $\mathscr{D}$ with respect to g. We have the direct sum decomposition $TM = \mathscr{D} \oplus \mathscr{D}^{\perp}$ and $P_{\mathscr{D}}$, $P_{\mathscr{D}^{\perp}}$ will denote the associated orthogonal projections on $\mathscr{D}$, $\mathscr{D}^{\perp}$ respectively.

DEFINITION 2.2. The Lagrangian function $L\colon TM \to \mathbb{R}$ of the system $(M, K, \mathscr{D}, V)$ is given by $L(w) = K(w) - V(\pi_{TM}(w))$, $w \in TM$.

2.1. *Some Hilbert Manifolds*

$\mathscr{H}^1(M)$ will denote the space of all curves $q\colon J \to M$, J an interval, which are absolutely continuous and the function $t \in J \mapsto K(\frac{Tq}{dt}(t))$ is locally integrable.

For $a_0, a_1 \in \mathbb{R}$, $a_0 < a_1$, let $H^1(M, [a_0, a_1])$ denote the subset of $\mathscr{H}^1(M)$ of all curves $q\colon [a_0, a_1] \to M$ contained in $\mathscr{H}^1(M)$. Given $m_0, m_1 \in M$, $H^1(M, [a_0, a_1], m_0)$ (resp. $H^1(M, [a_0, a_1], m_0, m_1)$) is the subset of $H^1(M, [a_0, a_1])$ of all curves q such that $q(a_0) = m_0$ (resp. $q(a_0) = m_0$, $q(a_1) = m_1$). Clearly $H^1(M, [a_0, a_1], m_0, m_1) \subset H^1(M, [a_0, a_1], m_0)$.

It is well known that $H^1(M, [a_0, a_1])$ is a Hilbert manifold and $H^1(M, [a_0, a_1], m_0)$, $H^1(M, [a_0, a_1], m_0, m_1)$ are submanifolds of it. If $q \in H^1(M, [a_0, a_1])$, the tangent space $T_q H^1(M, [a_0, a_1])$ to $H^1(M, [a_0, a_1])$ at q is the space of all H^1 sections η of the vector bundle $q^*TM \to [a_0, a_1]$ where q^*TM is the pull back of the tangent bundle $\pi_{TM}\colon TM \to M$ by q. This corresponds to the set of all H^1 curves $\eta\colon [a_0, a_1] \to TM$ such that $\pi_{TM} \circ \eta = q$. If $q \in H^1(M, [a_0, a_1], m_0)$ (resp. $H^1(M, [a_0, a_1], m_0, m_1)$), then $T_q H^1(M, [a_0, a_1], m_0)$ (resp. $T_q H^1(M, [a_0, a_1], m_0, m_1)$) is the subspace of all $\eta \in T_q H^1(M, [a_0, a_1])$ such that $\eta(a_0) = 0_{q(a_0)}$ (resp. $\eta(a_0) = 0_{q(a_0)}$, $\eta(a_1) = 0_{q(a_1)}$). Here 0_m, for $m \in M$, is the zero of the space $T_m M$. The manifold $H^1(M, [a_0, a_1])$ is endowed with the Riemannian metric G: if $\eta \in T_q H^1(M, [a_0, a_1])$, then $G(\eta) = \int_{a_0}^{a_1} g(\eta(t))\, dt$.

2.2. *Lagrangian Functional*

The Lagrangian function $L\colon TM \to \mathbb{R}$ defines a Lagrangian functional $\mathscr{L}\colon H^1(M, [a_0, a_1]) \to \mathbb{R}$ by $\mathscr{L}(q) = \int_{a_0}^{a_1} L(\frac{Tq}{dt})\, dt$. $\mathscr{L}$ is smooth.

2.3. *The $\mathscr{D}$-spaces*

Let us introduce the following subset of $H^1(M, [a_0, a_1])$:

$$H^1(M, \mathscr{D}, [a_0, a_1]) = \left\{ q \in H^1(M, [a_0, a_1]) \,\middle|\, \frac{Tq}{dt}(t) \in \mathscr{D}_{q(t)} \text{ for almost all } t \in [a_0, a_1] \right\}.$$

We define also:

$$H^1(M, \mathscr{D}, [a_0, a_1], m_0) = H^1(M, [a_0, a_1], m_0) \cap H^1(M, \mathscr{D}, [a_0, a_1]),$$

$$H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1) = H^1(M, [a_0, a_1], m_0, m_1) \cap H^1(M, \mathscr{D}, [a_0, a_1]).$$

Finally if $q \in H^1(M, [a_0, a_1])$ we set

$$H^1\mathscr{D}_q([a_0, a_1]) = \{\eta \in T_q H^1(M, [a_0, a_1], q(a_0), q(a_1)) \mid \eta(t) \in \mathscr{D}_{q(t)} \text{ for all } t \in [a_0, a_1]\}.$$

3. D'ALEMBERT VERSUS VAKONOMICS

We start with the definitions of d'Alembertian and vakonomic trajectories and after that we make a comparison between them.

3.1. *d'Alembertian Trajectories*

Let $q \in H^1(M, \mathscr{D}, [a_0, a_1])$. q is called a **d'Alembertian trajectory** of the mechanical system with constraints $(M, K, \mathscr{D}, V)$ if the differential $d\mathscr{L}(q)$ of $\mathscr{L}$ at q annihilates the subspace $H^1\mathscr{D}_q([a_0, a_1])$ of $T_q H^1(M, [a_0, a_1], q(a_0), q(a_1))$.

3.2. *Vakonomic Trajectories*

Let $q \in H^1(M, \mathscr{D}, [a_0, a_1])$. q is called a **vakonomic trajectory** of the mechanical system with constraints $(M, K, \mathscr{D}, V)$ if q is a stationary point for the restriction of $\mathscr{L}$ to the subset $H^1(M, \mathscr{D}, [a_0, a_1], q(a_0), q(a_1))$ of $H^1(M, [a_0, a_1], q(a_0), q(a_1))$. Note that this means: for any C^1 curve $\lambda \in]-\varepsilon, \varepsilon[\mapsto Q_\lambda \in H^1(M, [a_0, a_1], q(a_0), q(a_1))$, $\varepsilon > 0$, such that (i) $Q_0 = q$, (ii) $Q_\lambda \in H^1(M, \mathscr{D}, [a_0, a_1], q(a_0), q(a_1))$, then $\frac{d}{d\lambda}(\mathscr{L}(Q_\lambda))|_{\lambda=0} = 0$.

3.3. *Comparison of These Trajectories*

As we will see in Section 5, it may happen that $H^1(M, \mathscr{D}, [a_0, a_1], q(a_0), q(a_1))$ is not a submanifold of $H^1(M, [a_0, a_1], q(a_0), q(a_1))$. But if q is a smooth point of $H^1(M, \mathscr{D}, [a_0, a_1], q(a_0), q(a_1))$ then the tangent space $T_q H^1(M, \mathscr{D}, [a_0, a_1], q(a_0), q(a_1))$ of $H^1(M, \mathscr{D}, [a_0, a_1], q(a_0), q(a_1))$ at q is **not** $H^1\mathscr{D}_q([a_0, a_1])$ **unless** $\mathscr{D}$ is integrable. In that case $H^1(M, \mathscr{D}, [a_0, a_1], q(a_0), q(a_1))$ is always a submanifold and for any $q \in H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$ we have that

$$T_q H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1) = H^1\mathscr{D}_q([a_0, a_1]).$$

4. D'ALEMBERTIAN NON-HOLONOMIC MECHANICS

4.1. *The Total Second Fundamental Form of a Distribution*

We define a bilinear vector bundle morphism $B_{\mathscr{D}}: TM \times_M \mathscr{D} \to \mathscr{D}^\perp$ where $TM \times_M \mathscr{D}$ is the fiber product of the bundles TM and $\mathscr{D}$ and $\mathscr{D}^\perp$ is the distribution orthogonal complement of $\mathscr{D}$ with respect to g; let $(\bar{X}, \bar{Y}) \in T_m M \times \mathscr{D}_m$ and choose two germs of vector fields at m X, Y, so that $X(m) = \bar{X}$, $Y(m) = \bar{Y}$ and $Y \in \mathscr{D}$ (that means Y is a germ of section of $\mathscr{D}$). Then

$$B_{\mathscr{D}}(\bar{X}, \bar{Y}) := P_{\mathscr{D}^\perp}[(\nabla_X Y)(m)]. \tag{4.1}$$

This does not depend on the choice of X and Y.

LEMMA 4.1. *Let X, Y, P be three germs of vector fields of M at the point $m \in M$ such that $Y \in \mathscr{D}$ and $P \in \mathscr{D}^\perp$. Then the number $\langle \nabla_X P, Y \rangle(m)$ depends on the values $X(m)$, $Y(m)$, $P(m)$, only.*

Proof. If X', Y', P' are three other germs at $m \in M$ such that $X'(m) = X(m)$, $Y'(m) = Y(m)$, $P'(m) = P(m)$, $Y' \in \mathscr{D}$ and $P' \in \mathscr{D}^\perp$, then

$$\begin{aligned}
&\langle \nabla_{X'} P', Y' \rangle(m) - \langle \nabla_X P, Y \rangle(m) \\
&\quad = \langle \nabla_{X'-X} P', Y' \rangle(m) + \langle \nabla_X (P' - P), Y \rangle(m) + \langle \nabla_X P, Y' - Y \rangle(m).
\end{aligned}$$

Since $(X' - X)(m) = 0$ and $(Y' - Y)(m) = 0$, the first and third terms at the right hand side are zero. On the other hand, $P' - P = \sum f_j P_j$ where the f_j are germs of smooth functions vanishing at $m \in M$ and the P_j are germs at $m \in M$ of local vector fields P_j spanning $\mathscr{D}^\perp$. Then

$$\begin{aligned}
[\nabla_X (P' - P)](m) &= \sum (Xf_j)(m)\, P_j(m) + \sum f_j(m)(\nabla_X P_j)(m) \\
&= \sum (Xf_j)(m)\, P_j(m) \in \mathscr{D}^\perp.
\end{aligned}$$

So, $\langle \nabla_X (P' - P), Y \rangle(m) = 0$ because $Y \in \mathscr{D}$. ∎

Remark 4.1. If $\mathscr{D}$ is integrable, the restriction of $B_{\mathscr{D}}$ to $\mathscr{D} \times_M \mathscr{D}$ is just the classical second fundamental form of the leaves of the foliation defined by $\mathscr{D}$.

PROPOSITION 4.1. *A curve $q \in H^1(M, \mathscr{D}, [a_0, a_1])$ is a d'Alembertian trajectory if, and only if, it is smooth and satisfies the second order equation*

$$\nabla_t \frac{Tq}{dt} + P_{\mathscr{D}} \operatorname{grad} V(q_0) = B_{\mathscr{D}}\left(\frac{Tq}{dt}, \frac{Tq}{dt}\right) \tag{4.2}$$

(*here* $\nabla_t = \nabla_{\frac{Tq}{dt}}$).

COROLLARY 4.1. *The d'Alembertian trajectories are the projections on M of the trajectories of a spray.*

Proof of Proposition 4.1. Let $q \in H^1(M, \mathscr{D}, [a_0, a_1])$. We have the following expression for the differential

$$d\mathscr{L}(q) : T_q H^1(M, [a_0, a_1], q_0, q_1) \to \mathbb{R}:$$

if $\eta \in T_q H^1(M, [a_0, a_1], q_0, q_1)$, then

$$d\mathscr{L}(q)\,\eta = \int_{a_0}^{a_1} \left[\left\langle \frac{Tq(t)}{dt}, \nabla_t \eta(t) \right\rangle - \langle \operatorname{grad} V(q(t)), \eta(t) \rangle \right] dt.$$

q will be a d'Alembertian trajectory if, and only if, for all $\eta \in H^1\mathscr{D}_q$ $([a_0, a_1])$ one has

$$\int_{a_0}^{a_1} \left[\left\langle \frac{Tq(t)}{dt}, \nabla_t \eta(t) \right\rangle - \langle \operatorname{grad} V(q(t)), \eta(t) \rangle \right] dt = 0.$$

A classical reasoning shows that this implies that $q \in H^2(M, [a_0, a_1])$ (see Appendix I). Using integration by parts we get

$$\int_{a_0}^{a_1} \left[\left\langle \nabla_t \frac{Tq(t)}{dt}, \eta(t) \right\rangle + \langle \operatorname{grad} V(q(t)), \eta(t) \rangle \right] dt = 0,$$

for all $\eta \in H^1\mathscr{D}_q([a_0, a_1])$. This is equivalent to

$$P_{\mathscr{D}} \left[\nabla_t \frac{Tq}{dt} + \operatorname{grad} V(q) \right] = 0 \qquad \text{a.e on } [a_0, a_1].$$

By the definition of the total second fundamental form we know that

$$P_{\mathscr{D}^\perp} \nabla_t \frac{Tq}{dt} = B_{\mathscr{D}} \left(\frac{Tq}{dt}, \frac{Tq}{dt} \right) \qquad \text{a.e on } [a_0, a_1].$$

Hence we get

$$\nabla_t \frac{Tq}{dt} + P_{\mathscr{D}} \operatorname{grad} V(q) = B_{\mathscr{D}} \left(\frac{Tq}{dt}, \frac{Tq}{dt} \right) \qquad \text{a.e on } [a_0, a_1]. \tag{4.3}$$

In fact, Cauchy's theorem shows that all the solutions of (4.3) are smooth. ∎

Remark 4.2. Equation (4.2) shows that q satisfies the d'Alembert principle. In fact, the conservative external force along $q=q(t)$ is equal to $-\operatorname{grad} V(q(t))$ and so the difference between the acceleration $\nabla_t \frac{Tq}{dt}$ and the external force is equal to

$$B_{\mathscr{D}}\left(\frac{Tq}{dt}, \frac{Tq}{dt}\right) + P_{\mathscr{D}^\perp} \operatorname{grad} V(q).$$

4.2. *Invariant Volumes*

In order to construct a volume form on $\mathscr{D}$ we start, locally, choosing a special system of coordinates for TM. Let $(U, q^1, ..., q^n)$ be a local system of coordinates for M and $(\omega^1, ..., \omega^n)$ be an orthonormal basis of 1-forms. The sequence

$$(q^1 \circ \pi_{TM}, ..., q^n \circ \pi_{TM}, d_M q^1, ..., d_M q^n)$$

of real functions defined on $TU = (\pi_{TM})^{-1}(U)$ is a system of coordinates for the open set TU of TM; here $d_M q^i$ denotes the differential of the coordinate function $q^i\colon U \to \mathbb{R}$. Let us set $\hat{q}^i \overset{\text{def}}{=} d_M q^i\colon TU \to \mathbb{R}$, and write q^i instead of $q^i \circ \pi_{TM}$, $i = 1, ..., n$, for simplicity. The expressions of the 1-forms ω^i (also understood as functions on TU) are

$$\omega^i = \sum_{j=1}^{n} a^i_j d_M q^j = \sum_{j=1}^{n} a^i_j \hat{q}^j, \tag{4.4}$$

where $a^i_j\colon U \to \mathbb{R}$ are smooth functions; they induce the functions $\hat{\omega}^i\colon TU \to \mathbb{R}$

$$\hat{\omega}^i \overset{\text{def}}{=} \sum_{j=1}^{n} (a^i_j \circ \pi_{TM})\, \hat{q}^j, \qquad i = 1, 2, ..., n \tag{4.5}$$

and one can also consider new local coordinates for TM:

$$(TU; q^1, ..., q^n, \hat{\omega}^1, ..., \hat{\omega}^n). \tag{4.6}$$

The vector bundle $\mathscr{D}$ is an imbedded subbundle of TM and

$$\mathscr{D} \cap TU = \{v_p \in TU \mid \hat{\omega}^{m+1}(v_p) = \cdots = \hat{\omega}^n(v_p) = 0\}.$$

Then the restrictions of $q^1, ..., q^n, \hat{\omega}^1, ..., \hat{\omega}^m$ to the open set $\mathscr{D} \cap TU$ define a local system of coordinates for $\mathscr{D}$ on $\mathscr{D} \cap TU$; by consequence a volume form on $\mathscr{D} \cap TU$ is defined by the restriction $\tilde{\omega}$ to $\mathscr{D} \cap TU$ of the $(m+n)$-form

$$\omega = \omega^1 \wedge \cdots \wedge \omega^n \wedge d_{TM}\hat{\omega}^1 \wedge \cdots \wedge d_{TM}\hat{\omega}^m, \tag{4.7}$$

where each ω^i, in (4.7), means $\pi^*_{TM}\omega^i$, the pull-back of ω^i to TU; so, from (4.4)

$$\omega^i \cong \pi^*_{TM}\omega^i = \sum_{j=1}^{n} (a^i_j \circ \pi_{TM})\, d_{TM}(q^j \circ \pi_{TM}), \qquad i = 1, ..., n. \tag{4.8}$$

To obtain a (global) volume form on $\mathscr{D}$ one needs to assume that $\mathscr{D}$ is *orientable* as manifold. One way to obtain it is the following:

DEFINITION 4.1. A distribution $\mathscr{D}$ of constant rank m on a manifold M^n is orientable if there exists a differentiable exterior $(n-m)$-form Ψ on M such that, for any $q \in M$ and $z_1, ..., z_{n-m} \in \mathscr{D}^\perp_q$ then $\Psi_q(z_1, ..., z_{n-m}) \neq 0$ if, and only if, $(z_1, ..., z_{n-m})$ is basis of $\mathscr{D}^\perp_q$.

Remark that in the codimension one case $(m = n-1)$, $\mathscr{D}$ orientable is equivalent to the existence of a globally defined unitary vector field N, orthogonal to $\mathscr{D}_q$, $\forall q \in M$.

PROPOSITION 4.2. *The (local) volume form $\tilde{\omega}$ defined on $\mathscr{D} \cap TU$ by formula (4.7) is invariant under the flow of the vector field X induced by the fundamental Eq. (4.2) if, and only if, the trace of $B_{\mathscr{D}^\perp}/(\mathscr{D}^\perp \times_M \mathscr{D}^\perp)$ vanishes ($B_{\mathscr{D}^\perp}$ is the total second fundamental form of $\mathscr{D}^\perp$). If $\mathscr{D}$ is orientable, $\tilde{\omega}$ can be extended to a global volume on $\mathscr{D}$.*

Proof. Let $q_0 : [a_0, a_1] \to U \subset M$ be a solution of equation (4.2) and $z_0 = (Tq_0(t))/dt \cong (q_0, \dot{q}_0) : [a_0, a_1] \to TU$ its canonical lifting. Then $q^i(z_0) = q^i_0(t)$, $\hat{q}^i(z_0) = \dot{q}^i_0(t)$, $\hat{\omega}^\alpha(z_0) = \omega^\alpha(q_0; \dot{q}_0) = 0$, $\alpha = m+1, ..., n$ and

$$\frac{d\hat{\omega}^i}{dt}(z_0) = \frac{d\omega^i}{dt}(q_0; \dot{q}_0) = -\sum_{j=1}^{m} \omega^i_j(\dot{q}_0)\, \omega^j(\dot{q}_0), \qquad 1 \leqslant i \leqslant m \quad \text{(see [C])}.$$

We have $\omega^i_j = \sum_{k=1}^n A^i_{kj}\omega^k$ where $A^i_{kj} + A^j_{ki} = 0$, then

$$\begin{aligned}\frac{d\hat{\omega}^i(z_0)}{dt} &= -\sum_{j,k=1}^{m} A^i_{ki}(q_0)\, \omega^k(\dot{q}_0)\, \omega^j(\dot{q}_0)\\ &= -\sum_{j,k=1}^{m} A^i_{kj}(\pi_{TM}(z_0))\, \hat{\omega}^k(z_0)\, \hat{\omega}^j(z_0).\end{aligned}$$

Let us write simply A^i_{kj} instead of $A^i_{kj} \circ \pi_{TM}$; we know that the liftings above are the trajectories of the vector field X on $\mathscr{D}$ considered in the statement of the Proposition and is expressed in local coordinates $(q^1, ..., q^n, \hat{\omega}^1, ..., \hat{\omega}^n)$ of TU by

$$X = \sum_{i=1}^{n} \hat{q}^i \frac{\partial}{\partial q^i} - \sum_{i=1}^{m} \sum_{j,k=1}^{m} A^i_{kj}\hat{\omega}^k\hat{\omega}^j \frac{\partial}{\partial \hat{\omega}^i}. \tag{4.9}$$

We want to know in which conditions the volume form ω, given by

$$\omega=\omega^1 \wedge \cdots \wedge \omega^n \wedge d_{TM}\hat{\omega}^1 \wedge \cdots \wedge d_{TM}\hat{\omega}^m,$$

is invariant under X. Let us assume, for a moment, that $V=0$; we start by computing the Lie derivative $\Theta(X)\,\omega$:

$$\Theta(X)\,\omega=\sum_{i=1}^{n} \omega^1 \wedge \cdots \wedge \Theta(X)\,\omega^i \wedge \cdots \wedge \omega^n \wedge d_{TM}\hat{\omega}^1 \wedge \cdots \wedge d_{TM}\hat{\omega}^m$$
$$+\sum_{j=1}^{m} \omega^1 \wedge \cdots \wedge \omega^n \wedge d_{TM}\hat{\omega}^1 \wedge , ..., \wedge \Theta(X)\, d_{TM}\hat{\omega}^j \wedge \cdots$$
$$\wedge\, d_{TM}\hat{\omega}^m.$$

To obtain $\Theta(X)\,\omega$ is enough to compute $\Theta(X)\,\omega^i$ modulo

$$\omega^1, ..., \omega^{i-1}, \omega^{i+1}, ..., \omega^n, d_{TM}\hat{\omega}^1, ..., d_{TM}\hat{\omega}^m,$$

and, analogously, to compute $\Theta(X)\, d_{TM}\hat{\omega}^j$ modulo

$$d_{TM}q^1, ..., d_{TM}q^n, d_{TM}\hat{\omega}^1, ..., d_{TM}\hat{\omega}^{j-1}, d_{TM}\hat{\omega}^{j+1}, ..., d_{TM}\hat{\omega}^m.$$

But $\Theta(X)\,\omega^i=i(X)\, d_{TM}\omega^i+d_{TM}(\omega^i(X))=i(X)\, d_{TM}\omega^i+d_{TM}\hat{\omega}^i$, so

$$\Theta(X)\,\omega^i \equiv i(X)\, d_{TM}\omega^i, \qquad i=1, ..., n$$

(observe that on $\mathscr{D}$ we have $\hat{\omega}^\alpha=0$, $\alpha=m+1, ..., n$).

We also have $d_{TM}\omega^i=-\sum_{k=1}^n \omega^i_k \wedge \omega^k$ where, with the simplified notation, ω^i_k and ω^k mean $\tau^*_M\omega^i_k$ and $\tau^*_M\omega^k$, respectively. Thus

$$\Theta(X)\,\omega^i \equiv -\sum_{k=1}^{n} \omega^i_k(X)\,\omega^k+\sum_{k=1}^{n} \omega^k(X)\,\omega^i_k.$$

Since $\omega^i_k=\sum_{r=1}^n A^i_{rk}\omega^r$ and $A^i_{ri}=0$ $(\omega^i_i=0)$ for all r, we write

$$\Theta(X)\,\omega^i \equiv \sum_{j=1}^{m} A^i_{ij}\omega^j(X)\,\omega^i \equiv \sum_{j=1}^{m} A^i_{ij}\hat{\omega}^j\omega^i. \tag{4.10}$$

On the other hand $\Theta(X)\, d_{TM}\hat{\omega}^j=d_{TM}(\Theta(X)\,\hat{\omega}^j)$ and $\Theta(X)\,\hat{\omega}^j$ is the component of X along $\hat{\omega}^j$, that is, from (4.9) is equal to $-\sum_{k,\,l=1}^m A^j_{kl}\hat{\omega}^k\hat{\omega}^l$, $1\leqslant j\leqslant m$. Then

$$d_{TM}(\Theta(X)\,\hat{\omega}^j)=-\sum_{k,\,l=1}^{m} A^j_{kl}(\hat{\omega}^l d_{TM}\hat{\omega}^k+\hat{\omega}^k d_{TM}\hat{\omega}^l)-\sum_{k,\,l=1}^{m} \hat{\omega}^k\hat{\omega}^l d_{TM}A^j_{kl}.$$

But $d_{TM}A^j_{kl}\equiv 0$ modulo $d_{TM}q^1, ..., d_{TM}q^n$, so

$$\Theta(X)\, d_{TM}\hat{\omega}^j \equiv -\sum_{l=1}^{m} A^j_{jl}\hat{\omega}^l d_{TM}\hat{\omega}^j, \tag{4.11}$$

because $A^j_{kj}=0$. Finally, from (4.10) and (4.11) we have

$$\Theta(X)\,\omega=\left[\sum_{i=1}^{n}\sum_{j=1}^{m} A^i_{ij}\hat{\omega}^j-\sum_{j=1}^{m}\sum_{l=1}^{m} A^j_{jl}\hat{\omega}^l\right]\omega$$

and so

$$\Theta(X)\,\omega=\left[\sum_{\alpha=m+1}^{n}\sum_{j=1}^{m} A^\alpha_{\alpha_j}\hat{\omega}^j\right]\omega. \tag{4.12}$$

Equation (4.12) shows that $\Theta(X)\,\omega=0$ if, and only if,

$$\sum_{\alpha=m+1}^{n} A^\alpha_{\alpha j}=0, \qquad j=1, ..., m. \tag{4.13}$$

The intrinsic interpretation of conditions (4.13) comes from the consideration of the total second fundamental form $B_{\mathscr{D}^\perp}$ of the distribution $\mathscr{D}^\perp$. In fact $B_{\mathscr{D}^\perp}\colon TM\times_M \mathscr{D}^\perp\to\mathscr{D}$ is given by

$$B_{\mathscr{D}^\perp}(x, y)=\sum_{j=1}^{m} B^j(x, y)\, X_j$$

and

$$B^j(x, y)=-\sum_{\alpha=m+1}^{n}\omega^j_\alpha(x)\,\omega^\alpha(y)=\sum_{\alpha=m+1}^{n}\sum_{\beta=m+1}^{n} A^j_{\beta\alpha}\omega^\beta(x)\,\omega^\alpha(y).$$

But

$$\textit{trace of } B^j/(\mathscr{D}^\perp\times_M\mathscr{D}^\perp)=\sum_{\alpha=m+1}^{n} B^j(X_\alpha, X_\alpha)=\sum_{\alpha=m+1}^{n} A^j_{\alpha\alpha}=-\sum_{\alpha=m+1}^{n} A^\alpha_{\alpha j}. \tag{4.14}$$

Thus from (4.13) and (4.14) the proof of Proposition 4.2 follows for $V=0$. When $V\neq 0$ we have a similar proof because in that more general case, denoting the vector field by $X(V)$, one can write

$$X(V)=\sum_{i=1}^{n}\hat{q}^i\frac{\partial}{\partial q^i}-\sum_{j=1}^{m}\sum_{k,l=1}^{m} A^j_{kl}\hat{\omega}^k\hat{\omega}^l\frac{\partial}{\partial\hat{\omega}^j}-\sum_{j=1}^{m} V^j\frac{\partial}{\partial\hat{\omega}^j},$$

where $V^j = \omega^j(grad\ V)$, $j = 1, ..., m$. Then, clearly, $\Theta(X(V))\,\omega^i \equiv \Theta(X)\,\omega^i$, $i = 1, ..., n$, and $\Theta(X(V))\,\hat{\omega}^j = \Theta(X)\,\hat{\omega}^j + V^j$, $j = 1, ..., m$. So, $\Theta(X(V))\,d_{TM}\hat{\omega}^j = \Theta(X)\,d_{TM}\hat{\omega}^j + d_{TM}V^j$; but, we have $d_{TM}V^j \equiv 0$ modulo $\omega^1, ..., \omega^n$, because V depends only on q. Then $\Theta(X(V))\,\omega = \Theta(X)\,\omega$ and the proof is now complete. ▌

5. STUDY OF THE $\mathscr{D}$-SPACES

Next we shall study the vakonomic trajectories. For this we need a few facts about the Hilbert manifolds associated to distributions that we already called the $\mathscr{D}$-spaces.

5.1. *The Tangent Spaces of* $H^1(M, \mathscr{D}, [a_0, a_1], m_0)$

For the determination of the tangent structure to $H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ we need an explicit determination of it as a submanifold of $H^1(M, [a_0, a_1], m_0)$. To do this the most convenient way is to imbed the Riemannian manifold (M, g) isometrically into $(\mathbb{R}^N, \|\ \|)$ where $\|\ \|^2 = \sum_{i=1}^N dx_i^2$. This is possible with a suitable N, by the Nash–Moser imbedding theorem. For simplicity of notation we may assume $M \subset \mathbb{R}^N$ and, in this case, TM, $\mathscr{D}$ and $\mathscr{D}^\perp$ are subsets of $M \times \mathbb{R}^N$. Let E be the normal bundle over M, that is, the union $E = \bigcup_{m \in M} T_m^\perp M \subset M \times \mathbb{R}^N$ where $T_m^\perp M$ is the subset of $\mathbb{R}^N$ orthogonal to $T_m M$ with respect to the Riemannian manifold $(\mathbb{R}^N, \|\ \|)$. So we have the direct sum $T_m M \oplus T_m^\perp M = \mathbb{R}^N$ for each $m \in M$, and $\dim E = N$. Take now a tubular neighborhood (T, f) of M in $\mathbb{R}^N$ (see [L]) that means a smooth diffeomorphism $f\colon T \to \Omega$ from a open neighborhood T of the zero section in E onto an open set Ω in $\mathbb{R}^N$, $\Omega \supset M$, such that $f(0_m) = m$ for any zero vector $0_m \in E$, $m \in M$. If $\pi\colon M \times \mathbb{R}^N \to M$ is the first projection, the map $p = (\pi \mid E) \circ f^{-1}\colon \Omega \to M$ is a projection ($p^2 = p$); the pair (Ω, p) also represents the tubular neighborhood of M in $\mathbb{R}^N$. The set Ω is called the tube in $\mathbb{R}^N$ and T is said to be a tube in E; they play the same role and can be identified by the diffeomorphism f. The open set Ω, $M \subset \Omega \subset \mathbb{R}^N$, can be endowed with a distribution $\hat{\mathscr{D}}$ where $\hat{\mathscr{D}}_y$, $y \in \Omega$, is obtained from $\mathscr{D}_{p(y)} \subset T_{p(y)}M$ by translation (in $\mathbb{R}^N$). One can also define on Ω another distribution $\hat{\mathscr{D}}^\perp$ such that $\hat{\mathscr{D}}_y^\perp \subset \mathbb{R}^N$ is the orthogonal complement to $\hat{\mathscr{D}}_y$ with respect to $(\mathbb{R}^N, \|\ \|)$, that is, $\hat{\mathscr{D}}_y \oplus \hat{\mathscr{D}}_y^\perp = \mathbb{R}^N$. Denote by $P(y)\colon \mathbb{R}^N \to \hat{\mathscr{D}}_y^\perp$ the orthogonal projection. It is clear that $\hat{\mathscr{D}}_{|_M} = \mathscr{D}$ and that $\hat{\mathscr{D}}^\perp \cap TM = \mathscr{D}^\perp$. Given $q_0 \in H^1(\Omega, [a_0, a_1], m_0)$, the compacity of $q_0([a_0, a_1])$ implies that there exists a number $r > 0$ such that if $t \in [a_0, a_1]$ and $x \in \Omega$ are such that $\|q_0(t) - x\| < r$, then the restriction of $P(q_0(t))$ to $\hat{\mathscr{D}}_x^\perp \subset \mathbb{R}^N$ induces an isomorphism $\hat{\mathscr{D}}_x^\perp \to \hat{\mathscr{D}}_{q_0(t)}^\perp$. Let us denote by $H^1L^2(\hat{\mathscr{D}}^\perp, [a_0, a_1], m_0)$ the space of all equivalent classes of curves $(q, z)\colon [a_0, a_1] \to \hat{\mathscr{D}}^\perp$ such that $q\colon [a_0, a_1] \to \Omega$ belongs to $H^1(\Omega, [a_0, a_1],$

m_0), and that $t \in [a_0, a_1] \to \|z(t)\|$, $z(t) \in \hat{\mathscr{D}}^{\perp}_{q(t)} \subset \mathbb{R}^N$, is in L^2 (we also set that $z \in L^2(\hat{\mathscr{D}}^{\perp}_q, [a_0, a_1])$). Consider $U \subset H^1L^2(\hat{\mathscr{D}}^{\perp}, [a_0, a_1], m_0)$ as the subset of all classes (q, z) such that $\|q(t) - q_0(t)\| < r$ for all $t \in [a_0, a_1]$. Define

$$\Phi_U \colon U \to H^1(\mathbb{R}^N, [a_0, a_1], 0) \times L^2(\hat{\mathscr{D}}^{\perp}_{q_0}, [a_0, a_1])$$

as $\Phi_U(q, z) = (q', z')$ where $q' = q - q_0$ and for a.e $t \in [a_0, a_1]$, $z'(t) := P(q_0(t))\, z(t) \in \hat{\mathscr{D}}_{q_0(t)}$ (so $z' \in L^2(\hat{\mathscr{D}}^{\perp}_{q_0}, [a_0, a_1])$). Clearly the image of Φ_U in the Hilbert space $H^1(\mathbb{R}^N, [a_0, a_1], 0) \times L^2(\hat{\mathscr{D}}^{\perp}_{q_0}, [a_0, a_1])$ is an open subset and the Φ_U provide an atlas of charts of the manifold structure on $H^1L^2(\hat{\mathscr{D}}^{\perp}, [a_0, a_1], m_0)$. Define a mapping

$$\Pi \colon H^1(\Omega, [a_0, a_1], m_0) \to H^1L^2(\hat{\mathscr{D}}^{\perp}, [a_0, a_1], m_0)$$

as follows: if $q \in H^1(\Omega, [a_0, a_1], m_0)$, $\Pi(q)$ is the equivalence class of $(q, P(q)\,\frac{dq}{dt})$ where $z = P(q)\,\frac{dq}{dt}$ is the equivalence class of the curve $t \in [a_0, a_1] \mapsto z(t) = P(q(t))\,\frac{dq(t)}{dt}$. One can see that $H^1(\Omega, \hat{\mathscr{D}}, [a_0, a_1], m_0) = \Pi^{-1}(Z)$ where Z is the "zero section," that is, the submanifold of $H^1L^2(\hat{\mathscr{D}}^{\perp}, [a_0, a_1], m_0)$ defined as $Z = \{(q, 0_q) : q \in H^1(\Omega, [a_0, a_1], m_0)$, $0_q(t) = 0_{q(t)}\}$ where $0_{q(t)}$ is the zero of $\hat{\mathscr{D}}_q(t)^{\perp}$. For simplicity we set as j^1q the equivalence class of $(q, P(q)\,\frac{dq}{dt})$. Again let $q_0 \in H^1(\Omega, [a_0, a_1], m_0)$ and let

$$T_{q_0}\Pi \colon T_{q_0}H^1(\Omega, [a_0, a_1], m_0) \to T_{j^1q_0}H^1L^2(\hat{\mathscr{D}}^{\perp}, [a_0, a_1], m_0)$$

be the tangent mapping of Π at q_0. The chart (U, Φ_U) identifies $T_{j^1q_0}H^1L^2(\hat{\mathscr{D}}^{\perp}, [a_0, a_1], m_0)$ with the Hilbert space $H^1(\mathbb{R}^N, [a_0, a_1], 0) \times L^2(\hat{\mathscr{D}}^{\perp}, [a_0, a_1])$. Let $V_{q_0} \colon T_{q_0}H^1(\Omega, [a_0, a_1], m_0) \to L^2(\hat{\mathscr{D}}^{\perp}, [a_0, a_1])$ be the composition of $T_{q_0}\Pi$ with the canonical projection. Let us compute $V_{q_0}(\chi)$ for $\chi \in T_{q_0}H^1(\Omega, [a_0, a_1], m_0)$; take a C^1 curve $\lambda \in]-\varepsilon, \varepsilon[\to Q_\lambda \in H^1(\Omega, [a_0, a_1], m_0)$ such that $Q_0 = q_0$ and $(TQ_\lambda/d\lambda)_{|_\lambda} = 0 = \chi \in T_{q_0}H^1(\Omega, [a_0, a_1], m_0) \cong H^1(\mathbb{R}^N, [a_0, a_1], m_0)$. Then

$$V_{q_0}(\chi) = P(q_0)\,\frac{d\chi}{dt} + dP(q_0)[\chi]\,\frac{dq_0}{dt}.$$

Cauchy's theorem tells us that given $\eta \in L^2(\hat{\mathscr{D}}_{q0}, [a_0, a_1], m_0)$ there exists a χ such that

$$P(q_0)\,\frac{d\chi}{dt} + dP(q_0)[\chi]\,\frac{dq_0}{dt} = \eta$$

and $\chi(a_0) = 0$. This shows that V_{q_0} is surjective. Let, now, $q_0 \in H^1(\Omega, \hat{\mathscr{D}}, [a_0, a_1], m_0)$. Then $\Pi(q_0) \in Z$. But it is easy to see that the space normal to Z at $\Pi(q_0)$ is $L^2(\hat{\mathscr{D}}_{q0}{}^{\perp}, [a_0, a_1])$ in the identification of $T_{j^1 q_0} H^1 L^2(\hat{\mathscr{D}}^{\perp}_{q_0}, [a_0, a_1], m_0)$ with $H^1(\mathbb{R}^N, [a_0, a_1], 0) \times L^2(\hat{\mathscr{D}}^{\perp}_{q_0}, [a_0, a_1])$. Hence Π is transversal to Z. This shows that, since $m_0 \in M$, $H^1(\Omega, \hat{\mathscr{D}}, [a_0, a_1], m_0) = H^1(M, \hat{\mathscr{D}}_{|_M}, [a_0, a_1], m_0)$ (see Appendix II) is a submanifold of $H^1(\Omega, [a_0, a_1], m_0)$. If $q_0 \in H^1(M, \hat{\mathscr{D}}, [a_0, a_1], m_0) = H^1(\Omega, \hat{\mathscr{D}}, [a_0, a_1], m_0)$ and if $\lambda \in]-\varepsilon, \varepsilon[\to Q_\lambda \in H^1(\Omega, \hat{\mathscr{D}}, [a_0, a_1], m_0)$ is a C^1 curve such that $Q_0 = q_0$ and $(TQ_\lambda/d\lambda)_{|_\lambda} = 0 = \chi$, then $P(q_0) \frac{d\chi}{dt} = P_{\mathscr{D}^\perp} \nabla_t \chi$ and $dP(q_0)[\chi]\, dq_0/dt = -B_{\mathscr{D}}(\chi, Tq_0/dt)$, where as set, $P_{\mathscr{D}^\perp}$ is the orthogonal projection from the tangent bundle TM of M onto the subbundle $\mathscr{D}^\perp = \hat{\mathscr{D}}^\perp \cap TM$. Hence we get that $T_{q_0} H^1(M, \mathscr{D}, [a_0, a_1], m_0) = T_{q_0} H^1(\Omega, \hat{\mathscr{D}}, [a_0, a_1], m_0)$.

Proposition 5.1. *The $\mathscr{D}$-space $H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ is a submanifold of $H^1(M, [a_0, a_1], m_0)$ and the tangent space $T_{q_0} H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ at $q_0 \in H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ is the set of all $J \in T_{q_0} H^1(M, [a_0, a_1], m_0)$ (which is isomorphic to the H^1 sections of the pull back q^*TM of TM by q_0) such that $P_{\mathscr{D}^\perp} \nabla_t J = B_{\mathscr{D}}(J, Tq_0/dt)$.*

5.2. *The $\mathscr{D}$-space $H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$*

$H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$ is the subset of all curves $q \in H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ such that $q(a_1) = m_1$. Let us denote by $ev_1 \colon H^1(M, \mathscr{D}, [a_0, a_1], m_0) \to M$ the smooth map $ev_1(q) = q(a_1)$; it is clear that $ev_1^{-1}(m_1) = H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$. So $H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$ is closed in $H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ and we want to study the possibility of $ev_1^{-1}(m_1)$ to be a smooth submanifold of $H^1(M, \mathscr{D}, [a_0, a_1], m_0)$. For a given $q_0 \in H^1(\mathscr{D}, [a_0, a_1], m_0, m_1)$, q_0 is a **regular** point of ev_1 (which implies that $ev_1^{-1}(m_1)$ will be a submanifold in an open neighborhood of q_0 in $H^1(\mathscr{D}, [a_0, a_1], m_0)$) if, and only if, the derivative of ev_1 at q_0, $T_{q_0} ev_1 \colon T_{q_0} H^1(M, \mathscr{D}, [a_0, a_1], m_0) \to T_{q_0(a_1)} M$, is a surjection. If q_0 is not regular it is called a **critical point** of ev_1 and we say often that q_0 is a **singular curve** (see Remark 5.1, below). Then $T_{q_0} ev_1$ is not a surjection if, and only if, there exists a vector $w \neq 0$ in $T_{q_0(a_1)} M = T_{m_1} M$ such that $\langle J(a_1), w \rangle = 0$ for all $J \in T_{q_0} H^1(M, \mathscr{D}, [a_0, a_1], m_0)$. In order to analyze this condition we need some notation: $J = J' + J''$, $J' = P_{\mathscr{D}^\perp} J$, $J'' = P_{\mathscr{D}} J$; for $Y \in \mathscr{D}_m$, let us denote by $B_{\mathscr{D}}(Y) \colon T_m M \to \mathscr{D}^\perp_m$ the operator $B_{\mathscr{D}}(Y)\, X = B_{\mathscr{D}}(X, Y)$ and $B^*_{\mathscr{D}}(Y) \colon \mathscr{D}^\perp_m \to T_m M$ the adjoint of $B_{\mathscr{D}}(Y)$ with respect to $\langle\ ,\ \rangle$, that is, for any $\tilde{P} \in \mathscr{D}^\perp_m$ and any $X \in T_m M$, we have $\langle B^*_{\mathscr{D}}(Y)\, \tilde{P}, X \rangle = \langle \tilde{P}, B_{\mathscr{D}}(Y)\, X \rangle$. Call $B^{*\prime} = P_{\mathscr{D}^\perp} B^*_{\mathscr{D}}$, $B^{*\prime\prime} = P_{\mathscr{D}} B^*_{\mathscr{D}}$. We also have $w = w' + w''$, $w' \in \mathscr{D}^\perp_{m_1}$ and $w'' \in \mathscr{D}_{m_1}$. Let P be the vector field along $q_0 \in H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ with values on $\mathscr{D}^\perp$, solution of the Cauchy problem:

$$P_{\mathscr{D}^\perp} \nabla_t P + B^{*\prime}(\dot{q}_0)\, P = 0, \qquad P(a_1) = w'.$$

Then

$$\langle J(a_1), w' \rangle = \langle J(a_1), P(a_1) \rangle = \int_{a_0}^{a_1} [\langle \nabla_t J, P \rangle + \langle J, \nabla_t P \rangle]\, dt.$$

But

$$\langle \nabla_t J, P \rangle = \langle P_{\mathscr{D}^\perp}(\nabla_t J), P \rangle = \langle B_{\mathscr{D}}(J, \dot{q}_0), P \rangle = \langle J, B^*_{\mathscr{D}}(\dot{q}_0)\, P \rangle$$

and, so,

$$\langle J(a_1), P(a_1) \rangle = \int_{a_0}^{a_1} \langle J, \nabla_t P + B^*_{\mathscr{D}}(\dot{q}_0)\, P \rangle\, dt$$

$$= \int_{a_0}^{a_1} \langle J'', P_{\mathscr{D}} \nabla_t P + B^{*''}_{\mathscr{D}}(\dot{q}_0)\, P \rangle dt,$$

where $\dot{q}_0 = Tq_0/dt$. Since

$$0 = \langle J(a_1), w \rangle = \langle J(a_1), w' + w'' \rangle = \langle J(a_1), P(a_1) \rangle + \langle J''(a_1), w'' \rangle,$$

we have, for all $J \in T_{q_0} H^1(M, \mathscr{D}, [a_0, a_1], m_0)$:

$$0 = \langle J''(a_1), w'' \rangle + \int_{a_0}^{a_1} \langle J'', P_{\mathscr{D}} \nabla_t P + B^{*''}(\dot{q}_0)\, P \rangle\, dt. \tag{5.1}$$

But, the $J \in T_{q_0} H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ are characterized by $J(a_0) = 0$, $P_{\mathscr{D}^\perp} \nabla_t J = B_{\mathscr{D}}(J, \dot{q}_0)$ and this last equation can be written as

$$P_{\mathscr{D}^\perp} \nabla_t J' - B_{\mathscr{D}}(\dot{q}_0)\, J' + P_{\mathscr{D}^\perp} \nabla_t J'' - B_{\mathscr{D}}(\dot{q}_0)\, J'' = 0$$

that shows that J'' can be chosen arbitrarily such that $J''(a_0) = 0$ and J' is then solution of a Cauchy problem. Condition (5.1) above shows that

$$w'' = 0 \qquad \text{and} \qquad P_{\mathscr{D}} \nabla_t P + B^{*''}(\dot{q}_0)\, P = 0.$$

Finally, P is a vector field along q_0 with values on $\mathscr{D}^\perp$ such that

$$P(a_1) = w' \text{ and } \nabla_t P + B^*_{\mathscr{D}}(\dot{q}_0)\, P = 0.$$

Conversely, if there exists such a non-zero P, then for all $J \in T_{q_0} H^1(M, \mathscr{D}, [a_0, a_1], m_0)$ one has $\langle J(a_1), P(a_1) \rangle = 0$. Then $T_{q_0} ev_1$ is not a surjection. One can state:

PROPOSITION 5.2. *A curve $q_0 \in H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$ is a critical point of the evaluation map ev_1 if, and only if, there exists a non-zero vector field P along q_0 with values on $\mathscr{D}^{\perp}$ such that $\nabla_t P + B^*_{\mathscr{D}}(\dot{q}_0)\, P = 0$.*

Remark 5.1. The curves defined in Proposition 5.2 as critical points of the evaluation map ev_1 are the so called singular curves. One can show that they not depend on the metric $g = \langle\ ,\ \rangle$ but only on $\mathscr{D}$. To see this let us introduce the subbundle $\mathscr{D}^0$ of the cotangent bundle T^*M, annihilator of $\mathscr{D}$: for $m \in M$, $\mathscr{D}^0_m = \{z \in T^*_m M : z(v) = 0 \text{ for all } v \in \mathscr{D}_m\}$. $\mathscr{D}^0$ is a submanifold of T^*M of dimension $2n - m$ where m is the rank of $\mathscr{D}$. For each $z \in \mathscr{D}^0$, let K_z denote the subspace of the tangent space $T_z \mathscr{D}^0$ of $\mathscr{D}^0$ at z defined as the kernel of $\omega_0(z)$, ω_0 being the canonical symplectic 2–form on T^*M: $p \in K_z$ if for every $u \in T_z \mathscr{D}^0$ one has $\omega_0(z)(p, u) = 0$. A curve $q \in H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$ is singular (that is q is a critical point of ev_1) if, and only if, there exists a curve $z: [a_0, a_1] \to \mathscr{D}^0$, $q = \pi_{T^*M} \circ z$, such that for a.e $t \in [a_0, a_1]$, $Tz(t)/dt \in K_{z(t)}$.

6. THE VARIATIONAL NON-HOLONOMIC (VAKONOMIC) MECHANICS

Let q_0 be a vakonomic trajectory of a mechanical system with non holonomic constraints $(M, K, \mathscr{D}, V)$. Assume that q_0 is a regular point of ev_1 in $H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$. Then we have

$$\int_{a_0}^{a_1} [\langle \dot{q}_0, \nabla_t J \rangle - \langle grad\ V(q_0), J \rangle]\, dt = 0, \tag{6.1}$$

for all $J \in T_{q_0} H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$.

The set of all such J is ker D, the kernel of the operator D: $T_{q_0} H^1(M, [a_0, a_1], m_0, m_1) \to L^2(\mathscr{D}^{\perp}_{q_0}, [a_0, a_1])$, given by

$$DJ = P_{\mathscr{D}^{\perp}} \nabla_t J - B_{\mathscr{D}}(J, \dot{q}_0) = P_{\mathscr{D}^{\perp}} \nabla_t J - B_{\mathscr{D}}(\dot{q}_0)\, J.$$

The annihilator of ker D is the closure of the image of the adjoint D^* of D

$$D^*: L^2(\mathscr{D}^{\perp}_{q_0}, [a_0, a_1]) \to T_{q_0} H^1(M, [a_0, a_1], m_0, m_1)$$

relative to the inner products of L^2 and H^1; for $J, K \in T_{q_0} H^1(M, [a_0, a_1], m_0, m_1)$,

$$\langle J, K \rangle_{H^1} = \int_{a_0}^{a_1} \langle \nabla_t J, \nabla_t K \rangle\, dt.$$

Let us compute D^*: for $J \in T_{q_0} H^1(M, [a_0, a_1], m_0, m_1)$ and $P \in L^2(\mathscr{D}^\perp_{q_0}, [a_0, a_1])$ we have

$$\langle DJ, P\rangle_{L^2} = \langle J, D^*P\rangle_{H^1} = \int_{a_0}^{a_1} [\langle \nabla_t J, P\rangle - \langle B_{\mathscr{D}}(\dot{q}_0)\, J, P\rangle]\, dt$$

so

$$\langle DJ, P\rangle_{L^2} = \int_{a_0}^{a_1} [\langle \nabla_t J, P\rangle - \langle J, B^*_{\mathscr{D}}(\dot{q}_0)\, P\rangle]\, dt$$

$$= \int_{a_0}^{a_1} [\langle \nabla_t P - L(P)\rangle]\, dt,$$

where $L(P)$ is the unique vector field $L = L(P) \in T_{q_0} H^1(M, [a_0, a_1])$ such that:

(i) $\nabla_t L = B^*_{\mathscr{D}}(\dot{q}_0)\, P$.

(ii) The vector field $K_P \in T_{q_0} H^1(M, [a_0, a_1])$ such that $K_P(a_0) = O_{m_0}$ and $\nabla_t K_P = P - L(P)$ verifies $K_P(a_1) = O_{m_1}$ (see Remark 6.1 below).

We then have $\langle DJ, P\rangle_{L^2} = \int_{a_0}^{a_1} \langle \nabla_t J, \nabla_t K_P\rangle\, dt$ and $K_P \in T_{q_0} H^1(M, [a_0, a_1], m_0, m_1)$. Then $D^*P = K_P$. It is clear, from the definition of D^*, that the image $I_m D^*$ of D^* is closed.

Let us come back to q_0; relation (6.1) and the fact that $J(a_0) = O_{m_0}$ and $J(a_1) = O_{m_1}$, imply

$$0 = \int_{a_0}^{a_1} \langle \dot{q}_0 + W, \nabla_t J\rangle\, dt \qquad \text{for all} \quad J \in \ker D, \tag{6.2}$$

where W is the (unique) vector field of $T_{q_0} H^1(M, [a_0, a_1])$ such that

(i) $\nabla_t W = grad\, V \circ q_0$.

(ii) The vector field $U \in T_{q_0} H^1(M, [a_0, a_1])$ defined by $U(a_0) = O_{m_0}$ and $\nabla_t U = \dot{q}_0 + W$, verifies $U(a_1) = O_{m_1}$ (see Remark 6.1 below).

Since $\dot{q}_0 + W$ belongs to the annihilator of $\ker D$ and that this annihilator is $I_m D^*$, there exists a $P \in L^2(\mathscr{D}^\perp_{q_0}, , [a_0, a_1])$, such that $\dot{q}_0 + W = P - L(P)$ (in fact (6.2) can be written as $\langle U, J\rangle_{H^1} = 0$ for all $J \in \ker D$) and then there is such a P so that $U = D^*P = K_P$. Then $\nabla_t U = \nabla_t K_P$, that is, $\dot{q}_0 + W = P - L(P)$. This shows that $P \in H^1(M, \mathscr{D}^\perp, [a_0, a_1])$ and by covariant derivative: $\nabla_t \dot{q}_0 + \nabla_t W = \nabla_t P - \nabla_t L(P)$, or

$$\nabla_t \dot{q}_0 - \nabla_t P - B^*_{\mathscr{D}}(\dot{q}_0)\, P + grad\, V \circ q_0 = 0. \tag{6.3}$$

PROPOSITION 6.1. *A regular curve $q_0 \in H^1(M, \mathscr{D}, [a_0, a_1], m_0, m_1)$ is a vakonomic trajectory provided that there is a vector field $P \in H^1(M, \mathscr{D}^\perp, [a_0, a_1])$ such that*

$$\nabla_t \dot{q}_0 - \nabla_t P - B^*_{\mathscr{D}}(\dot{q}_0)\, P + grad\ V \circ q_0 = 0.$$

Moreover P is unique.

Remark 6.1.

(a) Let $J \in T_{q_0} H^1(M, [a_0, a_1], m_0, m_1)$ and $Q \in L^2(T_{q_0} M, [a_0, a_1])$. Then

$$\int_{a_0}^{a_1} \langle J, Q \rangle_g \, dt = -\int_{a_0}^{a_1} \langle \nabla_t J, K \rangle \, dt$$

where K is any element in $K \in T_{q_0} H^1(M, [a_0, a_1])$ such that $\nabla_t K = Q$. This follows from the fact that $J(a_0) = 0_{m_0}$ and $J(a_1) = 0_{m_1}$.

(b) Let J and Q as in (a). We want to represent $\int_{a_0}^{a_1} \langle \nabla_t J, Q \rangle \, dt$ as an inner product $\langle J, K \rangle_{H^1}$ where $K \in T_{q_0} H^1(M, [a_0, a_1], m_0, m_1)$. Since $\langle J, K \rangle_{H^1} = \int_{a_0}^{a_1} \langle \nabla_t J, \nabla_t K \rangle \, dt$, it is enough to find K such that $\nabla_t K = Q$ and $K(a_0) = 0_{m_0}$, $K(a_1) = 0_{m_1}$. That is in general impossible because the Cauchy problem $\nabla_t K = Q$, $K(a_0) = 0_{m_0}$, determines K. But observe that $J \in T_{q_0} H^1(M, [a_0, a_1], m_0, m_1)$ and that the orthogonal of this last space in $T_{q_0} H^1(M, [a_0, a_1])$ is the space

$$CL = \{ R \in T_{q_0} H^1(M, [a_0, a_1]) : R(t) = U_1(t) + t U_2(t) \}$$

(where U_1, $U_2 \in T_{q_0} H^1(M, [a_0, a_1])$ and $\nabla_t U_1 = \nabla_t U_2 = 0$). So, if $K_0 \in T_{q_0} H^1(M, [a_0, a_1])$ is such that $\langle J, K_0 \rangle_{H^1} = \int_{a_0}^{a_1} \langle \nabla_t J, Q \rangle \, dt$, then, for all $R \in CL$ we will have $\langle J, K_0 + R \rangle_{H^1} = \int_{a_0}^{a_1} \langle \nabla_t J, Q \rangle \, dt$. But one can choose U_1 and U_2 in R such that $(K_0 + R)(a_0) = 0_{m_0}$, $(K_0 + R)(a_1) = 0_{m_1}$, $U_1(a_0) = -K_0(a_0)$ and $U_1(a_1) = -(1/a_2 - a_1)[K_0(a_1) + U_1(a_1)]$.

(c) In order to determine $L(P)$ (resp. W) we choose first a $L_0 \in T_{q_0} H^1(M, [a_0, a_1])$ (resp. $W \in T_{q_0} H^1(M, [a_0, a_1])$) such that $\nabla_t L_0 = B^*_{\mathscr{D}} P$ (resp. $\nabla_t W = grad\ V \circ q_0$). Take $Q = P - L_0$ (resp. $Q = \dot{q}_0 + W$) in (b) above and construct $K_P \in T_{q_0} H^1(M, [a_0, a_1], m_0, m_1)$ (resp. $U \in T_{q_0} H^1 (M, [a_0, a_1],\ m_0, m_1)$) such that $\langle J, K_P \rangle_{H^1} = \int_{a_0}^{a_1} \langle \nabla_t J, P - L_0 \rangle \, dt$ (resp. $\langle J, U \rangle_{H^1} = \int_{a_0}^{a_1} \langle \nabla_t J, \dot{q}_0 + W \rangle \, dt$). Then take $L(P) = P - \nabla_t K_P$ (resp. $W = \nabla_t U - \dot{q}_0$).

Remark 6.2. One can see, from the last Proposition 6.1, that Eq. (6.3) induces on $TM = \mathscr{D} \times_M \mathscr{D}^\perp$ a flow whose trajectories are of the type $(\dot{q}(t), P(t))$. As we see the motions satisfying (6.3) that start at $(\dot{q}(0), P(0))$

$\in \mathscr{D} \times_M \mathscr{D}^\perp$ will be compatible with $\mathscr{D}$ (in the sense that $\dot{q}(t) \in \mathscr{D}_{q(t)}$ for all t) and also $P(t) \in \mathscr{D}^\perp_{q(t)}$ for all t.

The next result states that the above flow on $\mathscr{D} \times_M \mathscr{D}^\perp$ is, in fact, the flow of a Hamiltonian vector field.

PROPOSITION 6.2. *The equation $\nabla_t \dot{q} + \mathrm{grad} V \circ q = \nabla_t P + B^*_{\mathscr{D}}(\dot{q})\, P$ defines on T^*M a Hamiltonian vector field of Hamiltonian function $H\colon T^*M \to \mathbb{R}$, given by*

$$H(\alpha) = V(\pi_{T^*M}\alpha) + \tfrac{1}{2} \sup\{\alpha(v)/\langle v, v\rangle : v \in \mathscr{D} - \{0\}\}, \qquad \forall \alpha \in T^*M.$$

Proof. It is enough to consider the vector field X_V defined on $TM = \mathscr{D} \times_M \mathscr{D}^\perp$ by Eq. (6.3) and show that $\omega_0(\mu_* X_V, \cdot) = dH(\cdot)$ where ω_0 is the canonical symplectic form of T^*M and $\mu\colon TM \to T^*M$ is the diffeomorphism given by $\mu(v)(\cdot) = \langle v, \cdot \rangle$, for all $v \in TM$. ∎

APPENDIX I

We will prove that $q \in H^2(M, [a_0, a_1])$, needed in Proposition 4.1. The question is local. Take $t_0 \in [a_0, a_1]$ and choose a smooth orthonormal frame field $(e_1, ..., e_n)$, $n = \dim M$, in an open neighborhood U of $q(t_0)$ such that $(e_1, ..., e_m)$ is an orthonormal frame of $\mathscr{D}$ in U. Denote by $(\omega^1, ..., \omega^n)$ the corresponding coframe. Choosing $\eta \in H^1\mathscr{D}_q([a_0, a_1])$ with support in $]\, t_0 - \varepsilon, t_0 + \varepsilon[$ where $q(]\, t_0 - \varepsilon, t_0 + \varepsilon[) \subset U$ we have that

$$\int_{t_0-\varepsilon}^{t_0+\varepsilon} \sum_{i=1}^m \omega^i\left(\frac{Tq}{dt}\right) \omega^i(\nabla_t \eta)\, dt = \int_{t_0-\varepsilon}^{t_0+\varepsilon} \langle \mathrm{grad}\ V(q(t)), \eta(t) \rangle\, dt,$$

$\omega^i(\nabla_t \eta) = d\omega^i(\eta)/dt + \sum_{j=1}^m \omega^i_j(\frac{Tq}{dt})\, \omega^j(\eta)$, where ω^i_j are the Levi–Civita connection forms (note that $\omega^\alpha(\eta) = 0$ if $m+1 \leqslant \alpha \leqslant n$). Then for all $\eta \in H^1\mathscr{D}_q([a_0, a_1])$, $\mathrm{supp}\, \eta \subset]t_0 - \varepsilon, t_0 + \varepsilon[$ one has

$$\int_{t_0-\varepsilon}^{t_0+\varepsilon} \sum_{i=1}^m \omega^i\left(\frac{Tq}{dt}\right) \frac{d\omega^i(\eta)}{dt}\, dt$$

$$= \int_{t_0-\varepsilon}^{t_0+\varepsilon} \left[\sum_{i=1}^m \left[\omega^i(\mathrm{grad}\ V(q(t)) - \sum_{j=1}^m \omega^j\left(\frac{Tq}{dt}\right) \omega^j_i\left(\frac{Tq}{dt}\right)\right] \omega^i(\eta) \right] dt.$$

We can take for the $\omega^i(\eta)$, $1 \leqslant i \leqslant m$, arbitrary functions in $H^1(\mathbb{R})$ with support in $]t_0 - \varepsilon, t_0 + \varepsilon[$. Hence, last equality shows that the functions $\omega^i(\frac{Tq}{dt})$ are in $H^1(\mathbb{R})$, $1 \leqslant i \leqslant m$, and that

$$\frac{d}{dt}\left(\omega^i\left(\frac{Tq}{dt}\right)\right) = -\omega^i(\mathrm{grad}\ V(q(t))) + \sum_{j=1}^m \omega^j\left(\frac{Tq}{dt}\right) \omega^j_i\left(\frac{Tq}{dt}\right).$$

This states that for $1 \leqslant i \leqslant n$

$$\omega^i\left(\nabla_t \frac{Tq}{dt}\right) + \omega^i(\text{grad } V(q(t))) = 0$$

(recall that $\omega^i_j + \omega^j_i = 0$, for all $1 \leqslant i, j \leqslant n$).

APPENDIX II

Let Ω be an open subset of $\mathbb{R}^N$, endowed with a distribution $\hat{\mathscr{D}}$ with constant rank m. Let $M \subset \Omega$ be a closed submanifold such that for any point $y \in M$, $\hat{\mathscr{D}}_y \subset T_y M$. Then if $q \in H^1(\Omega, \hat{\mathscr{D}}, [a_0, a_1], m_0)$ and $m_0 \in M$, q has values on M. It is clear that $H^1(M, \hat{\mathscr{D}}_{|M}, [a_0, a_1], m_0)$ is contained in $H^1(\Omega, \hat{\mathscr{D}}, [a_0, a_1], m_0)$, so $H^1(M, \hat{\mathscr{D}}_{|M}, [a_0, a_1], m_0) = H^1(\Omega, \hat{\mathscr{D}}, [a_0, a_1], m_0)$. Let $T_M = \{t \in [a_0, a_1] : q(t) \in M\}$. Since M is closed and q is continuous, T_M is closed in $[a_0, a_1]$. T_M contains a_0 since $q(a_0) = m_0 \in M$. Assume that $T_M \neq [a_0, a_1]$. Let $\bar{t} = \inf\{t \in [a_0, a_1] : t \notin T_M\}$, $a_0 \leqslant \bar{t} < a_1$. We can choose a (curvilinear) chart of $\mathbb{R}^N$, $(O, x_1, ..., x_m, y_1, ..., y_s, z_1, ..., z_u)$, $m = \text{rank } \mathscr{D}$, $m + s = n = \dim M$, such that:

(i) $q(\bar{t}) \in O$, $x_i(q(\bar{t})) = 0$, $y_j(q(\bar{t})) = 0$, $z_k(q(\bar{t})) = 0$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant s$, $1 \leqslant k \leqslant u$;

(ii) $M \cap O = \{z_1 = \cdots = z_u = 0\}$;

(iii) $\hat{\mathscr{D}}_{q(\bar{t})} = \{dy_1 = \cdots = dy_s = dz_1 = \cdots = dz_u = 0\}$.

Restricting O, if necessary, we can assume that $\hat{\mathscr{D}} = \{dy = A dx, dz = B\, dx\}$, where A: $O \to \text{Mat}(s \times m)$ and B: $O \to \text{Mat}(u \times m)$ are two smooth matrix valued functions. The fact that $\hat{\mathscr{D}}_y \subset T_y M$ for all $y \in M$ can be expressed as $B(x, y, 0) = 0$ for all $(x, y, 0) \in O$. Restricting O again, there exist smooth matrix valued functions B_k: $O \to \text{Mat}(u \times m)$ such that $B = \sum_{k=1}^{u} z_k B_k$. Let Q be a closed ball centered at $q(\bar{t})$ of positive radius, such that $Q \subset O$. There exists $\varepsilon_0 > 0$ such that $q(t) \in Q$ if $t \in [\bar{t} - \varepsilon, \bar{t} + \varepsilon]$. Let $\xi(t) = x(q(t))$, $\eta(t) = y(q(t))$ and $z(t) = \eta(t)$ for $t \in [\bar{t} - \varepsilon, \bar{t} + \varepsilon]$. Then if $t \in [\bar{t} - \varepsilon, \bar{t} + \varepsilon]$, $z(t) = \int_{\bar{t}}^{t} [\sum_{k=1}^{u} z_k(\tau)\, B_k(\xi(\tau), \eta(\tau), z(\tau)) \frac{d\xi(\tau)}{d\tau}]\, d\tau$. (Since T_M is closed, $\bar{t} \in T_M$ and $z(\bar{t}) = 0$). Since Q is compact, there exists a constant $C > 0$ such that $\|B_k(p)\, v\|_u \leqslant C/\sqrt{u}\, \|v\|_m$ for all $1 \leqslant k \leqslant u$, $p \in Q$, $v \in \mathbb{R}^m$; here $\|\ \|_u$ (resp. $\|\ \|_m$) means the Euclidean norm in $\mathbb{R}^u$ (resp. $\mathbb{R}^m$). Hence on $[\bar{t} - \varepsilon_0, \bar{t} + \varepsilon_0]$ we have

$$\|z(t)\|_u \leqslant \left| \int_{\bar{t}}^{t} \left[\frac{C}{\sqrt{u}} \sum_{k=1}^{u} |z_k(\tau)| \left\| \frac{d\xi(\tau)}{d\tau} \right\|_m \right] d\tau \right|$$

$$\leqslant C \left| \int_{\bar{t}}^{t} \left[\|z(\tau)\|_u \left\| \frac{d\xi(\tau)}{d\tau} \right\|_m \right] d\tau \right|.$$

Let $\varepsilon = \min\{\varepsilon_0, \frac{1}{4}(C^2 \int_{\bar{t}}^{\bar{t}+\varepsilon_0} \|\frac{d\xi(\tau)}{d\tau}\|_u^2 \, d\tau)^{-1}\}$ and let $\mu = \sup\{\|z(t)\|_u, \bar{t} \leqslant t \leqslant \bar{t}+\varepsilon\}$. Therefore for all $t \in [\bar{t}, \bar{t}+\varepsilon]$,

$$\|z(t)\|_u \leqslant C\mu \sqrt{\varepsilon} \left(\int_{\bar{t}}^{\bar{t}+\varepsilon_0} \left\| \frac{d\xi(\tau)}{d\tau} \right\|_m^2 d\tau \right)^{1/2} \leqslant \frac{\mu}{2}$$

and so we get $\mu \leqslant \frac{\mu}{2}$. Hence $\mu = 0$ and $q(t) \in M$ for all $t \in [\bar{t}, \bar{t}+\varepsilon]$. This contradicts the definition of $\bar{t}$.

ACKNOWLEDGMENTS

The second author thanks Carlos E. Harle, Helena M. A. Castro, Esmeralda Dias and Marcelo H. Kobayashi for many helpful conversations on the subject.

REFERENCES

[AKN] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt, Mathematical Aspects of classical and celestial mechanics, *in* "Dynamical Systems III, Encyclopaedia of Mathematical Sciences" (V. I. Arnold, Ed.), Chap. I, pp. 1–286, 1988.

[C] E. Cartan, Sur la Représentation Géometrique des Systèmes Matériels Non-holonomes, *Att. Cong. Int. Matem.* **4** (1928), 253–261.

[CF] F. Cardin and M. Favretti, On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints, *J. Geom. Phys.* **18** (1996), 295–325.

[CKO] H. M. A. Castro, M. Kobayashi, and W. M. Oliva, Partially hyperbolic Σ-geodesic flows, *J. Differential Equations*, in press.

[FO] G. Fusco and W. M. Oliva, Dissipative systems with constraints, *J. Differential Equations* **63** (1986), 362–388.

[Ko] J. Koiller, Reduction of some classical nonholonomic systems with symmetry, *Arch. Rational Mech. Anal.* **118** (1992), 113–148.

[Ku] I. Kupka, Géométrie Sous–Riemanniene, Séminaire Bourbaki, 48ème année. 1995–96, n° 817, juin 1996.

[L] S. Lang, "Introduction aux Variétés Différentiables," Dunot, Paris, 1967.

[LM] A. D. Lewis and R. M. Murray, Variational principles for constrained systems; theory and experiment, *Internat. J. Non-Linear Mech.* **30** (1995), 793–815.

[MP] E. Massa and F. Pagani, Classical dynamics of nonholonomic systems: a geometric approach, *Ann. Inst. H. Poincaré, Phys. Théor.* **55** (1991), pp. 511–544.

[VG] A. M. Vershik and V. Ya. Gershkovich, Nonholonomic dynamical systems, geometry of distributions and variational problems, *in* "Dynamical Systems VII, Encyclopaedia of Mathematical Sciences" (V. I. Arnold and S. P. Novikov, Eds.), pp. 4–79, 1990.

[Z] G. Zampieri, Nonholonomic versus vakonomic dynamics, *J. Differential Equations* **163** (2000), 335–347.